TITLE: A Statistician's Place in Assessing the Likely Operational Performance of Army Weapons and Equipment

A STATISTICIAN'S PLACE IN ASSESSING THE LIKELY OPERATIONAL PERFORMANCE OF ARMY WEAPONS AND EQUIPMENT

E. S. Pearson
University College, London, England

THE BACKGROUND OF THIS PAPER. It has been a special honour to receive an invitation from the organising committee of this Conference to make the journey from England and to address you today. In thinking how I could best repay the compliment, it seemed to me that I should look for a subject in illustrating which I could draw on my own particular experiences, gained in working for the British armed services both during and since the second world war. From 1939 to 1946 I was attached with a number of members of the University College, London Statistics Department, to the British Ordnance Board. This is an organisation of some historic interest for I believe its foundation can be traced back to an appointment made in 1414, the year before the Battle of Agincourt! It is now concerned with certain aspects of the development and acceptance of weapons for both the Army, the Navy and the Air Force. Then, for some years after the war, I was a member of the Ordnance Board Anti-aircraft Lethality Committee and very recently I have been pulled back to be chairman of an advisory committee concerned with the general problem of assessment in connection with army weapons and equipment.

My main experience was with the subject which has been described as terminal ballistics and in particular with the lethal effectiveness of anti-aircraft fire. We were concerned also with field artillery fire and with the medium and small bombs of those days, in so far as fragmentation of the casing rather than blast played an important part in their effectiveness. It is of course true that the weapons and the army requirements of 15-20 years ago have been to a large extent out-dated, but if I make my main topic today a piece of historical recording, it is because I believe that a number of general principles and lessons emerge from such a study which are still relevant to the practice of experimentation and analysis in Army Research today.

It seemed to me that there were two advantages in taking illustrations from World War II experience. In the first place I could speak of matters about which I had the 'feel' from first hand knowledge and so perhaps could be more interesting as well as convincing in any arguments put forward. Secondly, it was easier to be factual without running into the danger of using classified material. What I shall try to do, therefore, is to give you first some account of the difficulties with which we were faced in the years 1939-45 in constructing a model which could be used to help determine how to improve the effectiveness of anti-aircraft fire. In describing this problem, it should be possible to indicate a number of lessons which are still relevant in a much wider field. There are also many points of difference which it will be instructive to emphasise.

THE STATISTICIAN'S PLACE. I should perhaps confess straight away that I shall say very little about statistics or about what is commonly thought of as the design of experiments. To this extent you may think that the leading phrase in the title of this paper is misleading, unless you interpret the words in the personal sense as referring to the statistician who is giving this address! But there is, I think, a point here which I should like to make. At the fourth of this series of Conferences, held in 1958, Dr. A. W. Kimball read a paper with the title: "Errors of the 3^{rd} kind in statistical consulting"; in this he discussed and illustrated the fault of giving a perfectly sound statistical answer to a problem which is not the real one needing solution.

Many of us are I think conscious of what might perhaps be called an error of a 4^{th} kind; that which the statistician makes when he allows his interest in the statistical elements of a problem and its potential for statistical elegance and sophistication to obscure what should be his prime objective, the solution of the real matter at issue. The fault is not so much that wrong statistical methods are used (Kimball's 3^{rd} kind of error) but that the situation does not justify the use of any refined statistical methods at all until the outstanding problem has been solved of obtaining data which are both relevant and reliable. The statistician, indeed, is called upon to be a scientist in the fullest sense of that term--to apply scientific method, not merely statistical techniques, to the job on hand.

When he has completed some piece of mathematical or arithmetical analysis, he needs to ask himself searchingly: does this answer make sense? I can recall, as no doubt some of you can too, war-time reports which appeared both in my country and in yours, containing a pretty piece of algebraic development or some standard analysis of variance, the conclusions from which obviously did not make sense. Perhaps such reports from youthful enthusiasts would never have appeared but for the inevitable shortage of experienced and critical supervision in rapidly expanding organisations. They are likely, however, to discourage the idea that mathematics or statistics were of value in problems of weapon development and testing, because the experienced non-statistical layman, the military or naval technical officer who had the feel of the problems, could see at once that the data would not bear the confident interpretation which was often placed on them.

Certainly in my own experience at the Ordnance Board it was the physical difficulty in securing meaningful experimental data which had always to be faced. There was very little opportunity for design as it is understood in agricultural or biological trials. There was no paramount function for the application of advanced statistics--we used to say that the only statistical tools which were needed were the normal distribution in 1, 2 and 3 dimensions, the Poisson and the bionomial. But it is true to say that the statistician's training, with the understanding which should follow of the meaning of variation and correlation, of randomness and probability, with its emphasis on the importance of adopting a critical outlook on assumptions--all this is likely to provide an excellent preparation for the kind of work we are discussing, but on one essential condition--that the training has been carried out in conjunction with practical application to data analysis. The trend in the teaching of mathematical statistics at our universities today is often increasingly away from any real application to data.

There is another point which I think is worth emphasising. One of the surest ways to cure the statistician from any tendency to over-sophistication is to arrange that he is present at experiments or trials, the data from which he is to use. In this respect we were lucky in England; we attended firing trials on the Shoeburyness Ranges, we were hot on the scene after bombs had been dropped on parked aircraft, trucks and wooden dummies in slit trenches on a special bombing range in the New Forest, and--as a wartime experience--we might happen to be present at a gun-site when German aircraft were the target. Under such conditions it is easier to come to grips with the meaning and limitations of data.

<u>THE ANTI-AIRCRAFT PROBLEM.</u> First let me try to put this problem into its setting of 20 or more years ago. As far as the Ordnance Board group was concerned, we had not to consider the problems of the deployment of guns, of the acquisition of targets, of the handling of mass attacks or other important tactical matters. These were questions for the Anti-aircraft Command and its Operational Research Section which was formed in the summer of 1940. Our work was closely related to the question of design, to understand more clearly the individual relationship between predictor, gun, shell, fuze and enemy target in order to advise what improvements were possible and likely to be worthwhile.

In this field of research where the terminal action in which one is interested may be taking place several thousand feet above ground, no overall experiment bringing in all the factors concerned is conceivable; the reasons for this are so obvious that I do not need to list them. As a consequence, it is absolutely essential to construct a mathematical model of the terminal engagement, and then to consider how the parameters of this model may best be estimated. As in so many other problems of military science, the model even if necessarily simplified, serves as an essential means of defining the relationships of the situation, showing how research investigation can be broken into separate pieces and emphasising at what points our lack of sure information is greatest and most hampering.

Let me now outline the problem and its solution in some detail, first describing the mathematical model and then discussing the three main headings under which gaps in knowledge had to be filled, namely:

(i) positioning errors (until the introduction of the proximity fuze in 1943-44 it was easy to combine the error of the time fuze with the predictor, gun-laying and ballistic errors);

(ii) fragmentation characteristics of the shell;

(iii) target vulnerability.

The difficulties which had to be overcome, largely through ignorance of physical properties in this hitherto unexplored field, are I think sufficiently instructive to be worth including as part of the story. Much the same problems were I know faced later on (building perhaps on our experience) in Section T of the Applied Physics Laboratory at Silver Spring and the associated Proving Ground near Albuquerque, where research and trial work was carried out for the U. S. Navy. I did not myself have any direct contact with U. S. Army investigations.

THE MATHEMATICAL MODEL. The first simplified model which was used involved:

(a) A three-dimensional normal distribution of positioning errors about the target, with a major axis along the shell trajectory and the standard errors in directions perpendicular to this axis equal, i.e., the density contours were taken to be ellipsoids with circular cross sections in planes perpendicular to the principal axis.

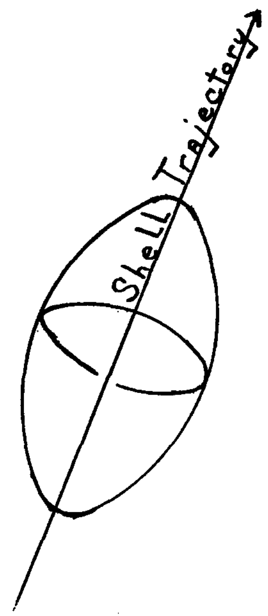

Figure 1

(b) A main fragment zone lying between two cones whose axis was that of the shell axis and the trajectory at time of burst, and a small sudsidiary nose cone.

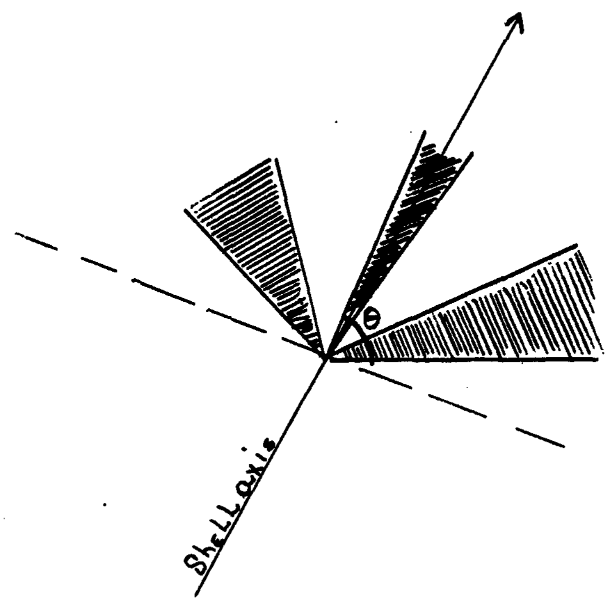

Figure 2

The density of fragments within the main zone was not of course uniform, though it might be treated as such for a first approximation. For any zone within which the average density of fragments of a given penetrating power could be regarded as constant, the probability distribution of strikes was taken as Poisson.

(c) For the aircraft, we first used what was termed an 'equivalent vulnerable target' represented by a sphere of a few feet in radius such that its 'perforation' by at least one 'lethal' shell fragment would result in a kill. Later, this representation had to be treated in more detail.

This simple model based on the trivariate normal and the Poisson distributions, with bounding surfaces consisting of ellipsoids, cones and spheres was amenable to computation, provided that meaningful numerical values for the various parameters could be estimated. But the task of filling in these unknown elements was immense and for a time the more we learnt, the more we realised our ignorance. Consider then some of the gaps to be filled.

<ins>THE POSITIONING ERRORS.</ins> The original data were collected from Practice Camp firings at towed 'sleeves', using kine-theodolites to measure the relative position of shell bursts and target. This was much too slow a target and the Practice Camp computational analysis was not very accurate. Later, in April 1940, a special trial of predictor accuracy was staged, following a free flying aircraft, and using camera recordings of the predictor output dials synchronised with kine-theodolites tracking the target. However, when German aircraft began to come over England later in 1940, it was at once clear that the aiming errors under operational conditions were much greater than those estimated from trials. We were up against the problem of increased operator inaccuracy under stress.

I remember P. M. S. Blackett (who was then in charge of the newly formed A. A. Command, Operational Research Group), wondering after watching the shell bursts in the night sky and a searchlight-held enemy aircraft, whether it would be possible to determine roughly an operational error distribution with appropriate photo-positioning equipment. I think that we later gave up all hope of estimating the actual aiming errors under operational conditions and made our calculations for a variety of different error combinations, which was often all that was needed in reaching conclusions about the relative merits of different types of shell, etc. It was only towards the end of the war when we were faced with that ideal straight-line-flying target, the V1 flying bomb, and when using proximity fuzes that a rough operational check on the overall adequacy of the model could be made.

THE FRAGMENTATION PROBLEM. Before the war, the standard trials for determining the fragmentation characteristics of shell were:

(a) Fragmentation in a sand-bag 'beehive', the shell fragments being recovered, passed successively through various sizes of sieve and (above a certain minimum size) counted and weighed.

(b) Trials to measure the dispersion and penetrating power of fragments by detonating the shell some 5 ft. above ground, in a surround of 2-inch-thick wooden targets, placed in a semicircle of, say, 30, 60, 90 or 120 ft. radius. The detonation was either at rest or obtained by firing the shell with appropriate remaining velocities against a light bursting screen.

With the war-time allocation of additional scientific effort onto weapon lethality problems, the number of questions which were posed for answering was greatly increased. The shell and bomb fragment attack on many targets besides aircraft had to be considered. On the one side it was necessary to have means of projecting individual fragments of various sizes at known velocities, against a variety of targets. On the other it was important to know more about the size-velocity-directional pattern as well as the retardation of the fragments projected by a complete shell burst in flight.

As soon as forward planning is attempted it becomes necessary to generalise the characteristics of a weapon; in the case of A. A. shell the ultimate objective was to be able to predict the characteristics of the fragment distribution from

(i) the drawing board design,

(ii) a knowledge of the particular explosive filling to be used, and

(iii) for any desired forward velocity of the shell.

It became clear that the old form of trials mentioned in the first paragraph of this section was inadequate. When shells were burst in flight in a wood target surround the resulting pattern of perforations could not be accurately related to the pattern from a static burst, merely by adding the component forward velocity of the shell. Nor was it easy to link the distribution of fragment sizes from the sand-bag collection with the number of perforations in the wood, using any simple assumptions about velocities and retardations. The essential need was for more basic physical experimentation; without this we could not generalise.

Here we were lucky in getting help from a very skilled scientific team at our Safety in Mines Research Establishment at Buxton, who initiated a programme of research which gradually succeeded in disentangling the picture. Shells on which small letters were engraved in successive rings round the circumference were fired at rest, within a surround of strawboard, against which a large number of small velocity measuring screens were placed. Fragments subsequently collected and weighed could be identified with a particular zone of the shell, and velocities estimated either by direct measurement or more crudely from depth of penetration into the strawboard.

It then became clear that the initial velocity of fragments varied very considerably with the part of the casing from which they came and similarly, that size or weight also varied with position. To some extent this initial velocity could be related to the charge/weight ratio of the section of the shell (perpendicular to its axis) from which the fragments originated. With this information, we began at last to get a surer picture of how fragments would be projected from different designs of shell detonated at any given velocity in free air.

It should be noted that the angle of the fragment zone, in particular the rather sharply defined 'cut-off angle' or semi-vertical angle θ of the backward bounding cone of my Figure 2 became particularly important with the introduction of proximity fuzes. If the pattern of fuze functioning was not co-ordinated with that of fragmentation the shell might generally burst in positions relative to the target such that fragments were bound to miss the more vulnerable parts of the aircraft.

<u>AIRCRAFT VULNERABILITY.</u> In the earliest trials carried out shortly before the war, an aircraft and an arc of large 2-inch thick vertical wooden screens were placed beyond and on opposite sides of a small burster screen at which the shell (with percussion fuze) was fired at a prescribed velocity.

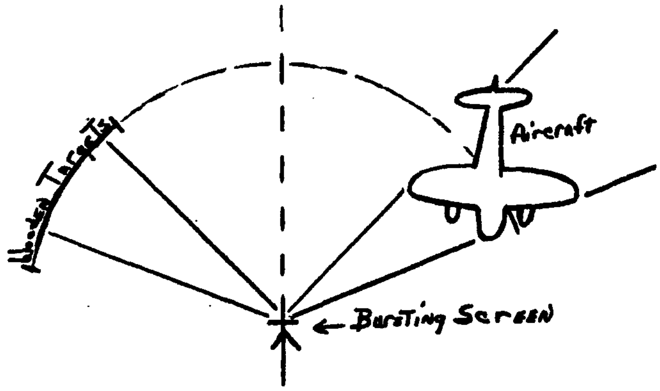

Figure 3

It was in this way possible to correlate the damage done to the aircraft with the density of fragments which perforated two inches of wood in a second, similarly constituted fragment stream. By noting and painting round the fragment holes after each round was fired, the same target could be used a large number of times, varying the aspect of attack and distance of detonation as desired.

It was from the observed correlation of density of 'throughs' (fragment capable of perforating 2 ins. of wood) and damage that it was possible to introduce into the model calculations a simplified 'equivalent vulnerable target'. This was the first method of attack. At a later stage after experimental techniques had become more refined and the Royal Aircraft Establishment assessors more experienced, it became possible to dissect the problem still further. The overall vulnerability picture was then built up from information gained by firing from high velocity barrels individual fragments of predetermined sizes, housed in specially designed cups, at a variety of aircraft components, which were screened where necessary by aluminium plates representing wing surfaces or fuselage.

The information so obtained could of course be used directly both in trying to draw conclusions about optimum fragment sizes and velocities and in considering ways of improving the protection of our own aircraft. Viewed in this way the problem may not appear to be statistical at all, but it did assume a statistical character as soon as one had to try and make use of this information in the 'model', with its shells bursting in a probability distribution around an aircraft and each projecting a composite stream of fragments, whose frequency distribution of strikes on equal areas of an intervening target would be roughly of Poisson form.

SOME CONCLUSIONS DRAWN FROM THIS SURVEY. Looking back now after a number of years, it seems to me that by 1944 we had really broken the back of the problem. It became possible to make recommendations with some confidence of a number of matters: on the optimum design characteristics of time fuzed and of proximity fuzed shell; on the relative importance of case thickness and explosive filling; on what might be achieved by using methods to control the size of fragments; on the relative gains to be won by improvement in fire control and in design of shell. Few such questions could have been answered with any confidence in 1939.

It is of course a truism that much of the fundamental research bearing on military problems is only rounded off when it is becoming too late to be of use in the war which provided the stimulus for the effort; and by the next war, the whole conditions of warfare are changed. This seems particularly true in regard to the ground-to-air weapons. But I think that the work I have been describing brought to the front a number of general principles, a sample of which I will bring to your attention in concluding this account.

The ease with which important factors may be overlooked. A common experience when the human mind starts to investigate the unknown is the way in which important considerations which seem so obvious afterwards are only realised through a process of slow and perhaps painful discovery.

(a) We did not for long appreciate the effect of ground ricochet in our firing trials. The influence of ricochet and other factors arising from proximity to the ground on the directional distribution of fragments would be natural operational effects in the case of field artillery or dropped bombs, but were very confusing when we were seeking information about the character of shell-bursts thousands of feet above ground. I know that the American experimenters appreciated this effect before we did and were the first to introduce ricochet traps into A. A. shell trials. Perhaps the most convincing demonstration of its existence which I recall occurred when we burst a 500 lb. bomb statically, with axis inclined at 30° to the vertical. The target screens showed a striking pattern of holes; a tilted belt like the forward-arm of a V from direct hits and another, like the other arm, from the ground ricochets. As long as bombs or shell were burst with their axes horizontal (or vertical), the effect remained unnoticed.

(b) Again, when studying the size distribution of fragments, the amount of secondary break-up on striking the collecting medium after detonation, was only realised when strawboard was used in place of sand

and the paths of these pieces, broken on first strike, could be traced through the successive layers of board.

(c) Another point not fully appreciated was the effect of emotional stress on the human element under battle conditions. The assessment of its magnitude, especially under circumstances and conditions which cannot be precisely foretold, is one of the hardest problems of the moment.

<u>The place of basic research</u>. In many instances it may not be too difficult to carry out a realistic trial of a particular weapon, against a given target under specified environmental conditions. But a more fundamental knowledge is necessary to assess the performance of weapons, perhaps still on the drawing board, under a wide variety of conditions. It was in this connection that the detailed experimental work on fragmentation performed to laboratory standards was essential, even if the laws of initial velocity, of size distribution and of retardation which resulted were to some extent empirical.

<u>The value of having something up your sleeve</u>. Observation of the amount of the metal casing which appeared to be broken up into dust or very small fragments*, on detonation, suggested that the destructive power of the anti-aircraft shell might be considerably increased by 'controlling' the size of fragments. It was over this matter that the help of the Safety in Mines Research Establishment was first called on, and by the end of the war this research group had developed a variety of techniques, relatively easily applied, by which it was possible to control the size and shape of shell and bomb fragments to a remarkable degree. These techniques were never used** but they were available to put into operation should any new target have had to be faced, e.g. a tough one against which only large fragments could be effective.

---

* It was realised later that some of this effect was due to secondary break-up of the large fragments on striking the collecting medium.

** It was found later that the Germans had applied a system of external grooving to some of their A. A. shell, apparently to increase the fragment size.

These are some of the still relevant points which I have noted in again coming into contact with problems of weapon research and development after a gap of several years. I am sure there are other lessons to be drawn from these World War II investigations, and without doubt those scientists who have carried on continuously in government service will have quietly absorbed them, so that they form part of their whole attitude of approach to the problems of today.

THE POSITION TODAY. There are, of course, many obvious differences between:

(a) The war-time problem, which was essentially that of trying to establish an understanding of a weapon system in service, in order to determine how its effectiveness could be improved, under conditions which were not expected to change radically from those known to exist; and

(b) The problem of today, which is greatly concerned with predictive assessments of the operational performance of future systems, taking many years to develop and to be used against an opponent whose future equipment, weapons and tactics must be to a large extent a matter of guesswork.

In the course of war, even when action has to be taken to meet a new situation, this can be done by working on the basis of information which possesses some element of reality. A good example of this occurred in 1944 with the launching of the V.1 flying bombs against London. Within a few days a complete bomb which had been shot down without exploding was recovered, and immediate steps could be taken to estimate its vulnerability to shell-fire and fighter attack.

As far as I can recall, priority trials were undertaken to determine (a) the burst pattern of a proximity fuze around such a target, and (b) the nature and extent of its vulnerability to A. A. shell fragments. How quickly we went as far as inserting these new parameters into our probability model, I cannot remember; but it must have been soon evident that the V.1 was a target which could be successfully engaged by 3.7 inch anti-aircraft guns with existing shell, provided they were supplied with proximity fuzes. The large-scale delivery of American fuzes and the appropriate re-deployment of guns, when achieved after some weeks when the fighter aircraft had been forced to take the leading defence roll, played a very large part in countering the menace.

The scientific effort, when it became accepted as of value by the armed services, was quite naturally first directed to the study of the performance of individual weapons or pieces of equipment; the radar set, the proximity fuze, the terminal ballistics of a shell or of a variety of anti-tank weapons. Today there is a special demand for scientific aid in the intractable job of peering into the future. The lead for this activity was of course provided by the Operational Research Sections which were closely associated with various operational commands during the war. In this very difficult field of prediction in which the last war's operational experience becomes less and less relevant, the scientific line of attack must consist in welding together a great number of elements.

The following scheme of relationships illustrates what I mean by the many-sided approach:

Mathematical models Paper study	$\longrightarrow$ $\longleftarrow$	Laboratory experiments Simulator trials Range trials
War games Computer games	$\longrightarrow$ $\longleftarrow$	Field experiments Army exercises

The overall inferences to be drawn from the whole build-up are not of course matters of statistics; but the use of the theory of probability and of stochastic processes is implicit in the studies on the left-hand column, while statistical planning plays its part in the laboratory experiments and the range trials--even to some extend in the field trials.

I have already tried to illustrate the great value of a mathematical model in forming the structure against which an evaluation problem may be broken up into parts for separate study. In so doing attention is drawn to the links in the construction where essential information in quantitative form is most needed and perhaps most lacking. Again, and this is important, by permitting a good deal of elasticity in the mechanism and allowing for the introduction of factors which might conceivably operate in a future situation, the model may be used to extrapolate beyond the envelope of engagement conditions tested during field trials or even accepted as likely under present combat conditions.

The application of the model approach to the problem of ground-to-air missile evaluation is the natural successor to the war-time investigations which I have described. The break-up of the problem for study under four headings still remains as before.

(a) Engagement geometry, (b) fuze performance,

(c) warhead effectiveness, (d) target vulnerability.

But problem (a) has taken a much more complex shape, involving perhaps the use of both analogue and digital computers. The war game has an essential part to play as a research tool in the combined attack on the problem of developing weapons, equipment and tactics for the future. Its main function is perhaps to aid thought and analysis rather than to obtain directo results. By injecting the human decision process into the study, it provides an insight into the complex nature of land battle which it would be hard to get in any other way. In this form of study, as elsewhere, the essential need to formulate rules, focuses attention on the limiting conditions which have to be accepted by whatever route we try to make predictions of the performance of future systems.

As a final illustration of where we now stand, let me refer to a problem of considerable present interest in whose solution a number of the techniques tabled above might be called in. This is the problem of comparing the merits of the <u>free flight gun</u> and the <u>guided missile</u> in the ground attack on armour. Both types of weapon depend, though in very different degree, on the human operator:

<u>The free flight gun.</u> Here we have a system, fairly well understood which has been studied for years and for which a reasonable idea of performance under operational conditions is available. The operator has only to concentrate while laying the gun and, after firing, plays no further part in the fate of that particular round. The greatest element of uncertainty lies in the vulnerability of his own gun, and to assess this requires rather extensive study of visibility and audibility in a variety of environments.

<u>The guided weapon.</u> The advantage of this weapon is that its firing position can be concealed behind the crest of a hill. However, the human controller who must see the target, has to concentrate for a considerable time (depending on the range) in guiding the weapon onto the target. That he can do this with fair success has been demonstrated on a simulator and with live weapons used under trial conditions. The open question here is whether he can maintain this performance in an operational setting, when subject to the fears and emotions to which he would be exposed in battle.

A sound basis for any policy decision on these alternative systems must depend on a comparative quantitative assessment; this cannot be completed without these missing pieces of information--the vulnerability of the gun to enemy counter action, the fall-off in human performance in a battle setting, and now adding to the puzzle, the observational power of the helicopter. Success in solution depends not only in not overlooking these considerations but, in persuading authority to provide the means of proceeding to the answers. How often one wonders have important decisions on weapon policy had to be taken in the past when the basic information for a real comparison was not available, although with greater foresight, perhaps, it might have been obtained in time.

Finally, it may again be asked: what of the statistician? Have I pushed him out of the picture: I think not. You must remember that I have been concentrating on a particular aspect of this matter of research, development and testing--the assessment of operational performance of weapons. In this peculairly difficult field, the statistician becomes the scientist who must merge his statistical identity into that of a group of men trained in several disciplines, but prepared to give no undue weight to any one of them in searching for answers to the problems in hand. That at any rate has been my personal experience.